PUPIL DIAMETER AND THE CROSS-ADAPTIVE CRITICAL TRACKING TASK; A METHOD OF WORKLOAD MEASUREMENT

Thomas Edward McFeely

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

PUPIL DIAMETER AND THE CROSS-ADAPTIVE
CRITICAL TRACKING TASK;
A METHOD OF WORKLOAD MEASUREMENT

by

Thomas Edward McFeely

Thesis Advisor: R. A. Hess

June 1972

Approved for public release; distribution unlimited.

Pupil Diameter and the Cross-Adaptive Critical

Tracking Task; a Method of Workload Measurement

by

Thomas Edward McFeely
Lieutenant, United States Navy
M.S., Naval Postgraduate School, 1971

Submitted in partial fulfillment of the
requirements for the degree of

AERONAUTICAL ENGINEER

ABSTRACT

Two new applications of established techniques for measuring an individual's level of stress (workload) in tracking tasks are presented. An indirect technique of measuring "reserve capacity" is utilized in a two-axis cross-coupled compensatory tracking task. A direct psychophysiological measurement is made by recording time histories of operator pupil diameter.

Results obtained indicate that each method yields a good index of workload, although considerable variance in the data is observed. The level of instability in the second axis of the cross-adaptive method is shown to be related to the level of workload in the primary axis. Increased pupil diameter is shown to be similarly related to operator workload. The simultaneous application of both techniques is determined to be inappropriate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

D - pupil diameter

D_o - initial unloaded pupil diameter

D_{sm} - maximum (peak) pupil diameter

D_{ss} - steady state pupil diameter

e_f - criterion error

e_p - primary task displayed error (vertical axis)

e_s - secondary task displayed error (horizontal axis)

e_T - total error ($|e_p| + |e_s|$)

K - gain

K_p - operator gain

P_p - primary task performance

s - Laplace operator

t - time

t_T - total run time

Δt_s - time interval of constant λ_{ss}

w - weighting function

$Y_p(s)$ - operator transfer function

λ - value of unstable root of first-order tracking task; level of instability

λ_c - critical, or limiting value of instability

λ_p - instability level of primary task

λ_{sm} - maximum (peak) instability level of secondary task

λ_{ss} - steady state instability level of secondary task

σ - standard deviation

τ_e - effective time delay

$(\bar{\ })$ - average or mean value

$(\ )^*$ - normalized value

$(\ )_w$ - weighted value

$|\ |$ - absolute value

ACKNOWLEDGEMENT

The author expresses his gratitude to the following individuals whose assistance and guidance made this thesis possible:

D. T. McRuer, president of Systems Technology, Inc., Hawthorne, California, for his cooperation and personal interest in the author's research during a six-week industrial tour with his company.

R. W. Allen, STI project engineer, who is principally responsible for the formulation of the critical cross-adaptive task and with whom the author collaborated on this research, for his selfless assistance and cooperation.

Assoc. Professor G. K. Poock, for the use of the Human Engineering Laboratory and its resources and particularly for the use of the pupil measuring equipment.

Lcdr. C. M. Bohley and Lt. H. G. Chalkley, fellow students, for donating a considerable amount of their time as subjects for the experiment.

Asst. Professor R. A. Hess, the author's thesis advisor and principal subject, for his time, unbounded patience and continuous support throughout this research.

Asst. Professor M. H. Redlin, the second reader of this thesis, for his comments, advice and constructive criticism.

His family, particularly his wife, for their cheerful and much-needed support during the long and difficult gestation of the thesis.

I. INTRODUCTION

As the performance, complexity and cost of modern aircraft weapons systems continue to accelerate, the physical and mental demands imposed upon their pilots become unmanageable. To operate within the narrowing margins of safety and effectiveness, pilots can no longer tolerate large attentional demands.

The development of boosted hydraulic control systems has reduced the physical loads on the pilot to a minimum, but the aggregate of mental loads, although reduced significantly in recent years, has hardly reached its nadir. The level of research activity in this area of workload reduction continues to remain high and the development of related equipment has progressed from simple autopilots to complex inertial guidance and navigation systems, from gyro horizons to situation indicators and flight directors, from the well-known "joy stick" to the new side-mounted hand controller. These devices have all been designed to increase the pilot's performance in the modern arena of the air by reducing his attentional demands; i.e., his mental workload.

The term workload refers here to a level of stress or the amount of effort expended to achieve a certain level of performance. Workload and performance are thus inter-connected but distinct. Kelley [Ref. 1] has shown that performance measurements are not in themselves valid workload indicators and that under certain circumstances increased performance

can be achieved at higher levels of performance-measured workload, and vice versa. Reference 2 similarly demonstrates this fact during an evaluation of helicopter autopilots.

This important distinction between workload and performance can be exemplified by the case of two jet aircraft performing a night carrier landing. One of the aircraft is equipped with an approach power compensator (an automatic throttle), and the other is not. Now consider the pilots. Both must achieve the same high level of performance to bring the aircraft aboard but their levels of workload are considerably different.

Thus, in order to properly evaluate an aircraft system or the effect of changes within that system, both operator performance and workload measurements must be taken.

Methods of measuring performance are now quite standard and commonly use some function of "system error" as the criterion. Methods of determining workload, however, have yet to be completely developed. There exists no "meter" which when plugged into the complex man-machine system gives direct readings of workload. The various techniques which have been employed in past research can be grouped into three categories:

1. subjective evaluations (pilot ratings)
2. measurement of "reserve capacity"
3. measurement of psychophysiological responses

Singly and in combination, these techniques have been shown to be relatively reliable indicators of workload.

Pilot ratings, when properly used, have been regarded as the most effective of these techniques and have been used to correlate or validate the other workload measures [Refs. 3,4]. By definition, however, these ratings are operator-variant and require a certain level of experience upon which to base the evaluation. In view of these requirements and in light of previous validation of the methods employed in this research by pilot rating [Refs. 4,5], this technique was not investigated.

Thus, in an attempt to provide a simple, operator-invariant method of workload measurement, this thesis has made use of techniques (2) and (3), above. The application of a cross-adaptive unstable tracking task was used to measure reserve capacity and the recording of changes in pupil diameter was used as the psychophysiological measure. Both methods were examined while subjects performed a compensatory tracking task.

II. THEORY

A. RESERVE CAPACITY

In the field of manual control the term "reserve capacity" (or "excess control capacity") is used to indicate the difference between an operator's maximum control capacity and that amount taken up by the particular task he is performing. The normalized ratio of this absorbed amount to the total amount of control is the attentional demand of the task, or simply the level of workload the task places on the operator. Thus by measuring this reserve capacity one is led to the "level of workload":

$$\text{Level of Workload} = \frac{\text{attentional demand}}{\text{total capacity}} = 1 - \frac{\text{reserve capacity}}{\text{total capacity}} \tag{1}$$

Note that the maximum level of workload, as defined here, is unity.

1. Operator Loading Tasks

The most common method of determining the reserve capacity is to load an operator with an additional, secondary task. This task may take several forms; a separate tracking task, a mental task (e.g. solving mathematical problems), or a responsive task (e.g. responding to light signals by pushing the correct button).

The reserve capacity of the operator can be measured in several ways. In the method most frequently used, the operator's performance on the primary (original) task is

determined, from this a criterion is established, and the operator is told to maintain this level of performance throughout the task. The level of difficulty of the secondary (loading) task is increased until the operator reaches the point where he can no longer meet the primary task performance criterion. This level of secondary task difficulty then becomes the measure of the operator's reserve capacity.

Knowles [Ref. 6] provides an excellent summary of this loading rationale and of early work in the area.

2. Cross-Adaptive Systems

A problem frequently encountered in using these secondary loading techniques is the difficulty of constraining the primary task performance. Although instructed otherwise, an operator may divert his attention from the primary task and focus on the secondary, particularly when the secondary level of difficulty is high. Thus his primary task performance deteriorates and the resulting higher level of difficulty he can tolerate in the secondary task is no longer a valid measure of his reserve capacity.

The so-called cross-adaptive technique overcomes this problem by automatically adjusting the level of difficulty on the secondary task on the basis of primary task performance. Although the operator may become distracted by a very difficult secondary task causing a lower level of performance in the primary, the adaptive system will sense the poorer performance and reduce the level of difficulty. The distraction is thus reduced and the primary task performance returns to

the desired level. In this way the operator is constrained to maintain the prescribed level of primary performance and an accurate measure of his reserve capacity can then be made.

A schematic of this type system is shown in Fig. (1).

Kelley [Ref. 7] points out that there are three elements of such an adaptive system. They are: (1) a method of measuring performance, (2) an adjustable parameter which changes the level of difficulty (called the adaptive variable) and (3) the adaptive logic which automatically changes the adaptive variable as a function of the performance measurement.

The specific application of these elements to this thesis is described in Section III.

3. Critical Instability Task

In a report which is generally regarded as a classic in the field of control theory, McRuer et al. [Ref. 8] presented a complex mathematical formulation of a human operator in a compensatory tracking task. This quasi-linear model consists of a linear describing function and the remaining non-linearities which are grouped and referred to as "remnant." This remnant is thought to be caused by non-linear or unsteady operator behavior and by the superposition of noise on the operator's linear output.

A simplified model of the human describing function, termed the "cross-over" model, was found to give a fairly good representation of the human operator under certain circumstances. It consisted principally of a pure operator gain (K_p) and an effective time delay (τ_e). This τ_e term is shown

to be a random variable consisting of various operator delays, lags and high frequency leads.

Jex et al. [Ref. 9] proved that when an operator is controlling a first-order divergent element with transfer function $Y_c(s) = \frac{\lambda}{s-\lambda}$, the maximum value of λ under which control can be maintained is approximately equal to the inverse of the operator's effective time delay. This instability level is termed "critical" (λ_c) and provides a simple and direct measure of an operator's tracking ability.

The task in which λ is slowly and monotonically increased during a run until control is lost is called the "critical task." A similar task in which λ is maintained at some controllable value is called the "sub-critical task."

Applications of this critical task appear limitless in the area of human response research. It is currently employed by a number of experimenters as an indicator of psychomotor performance. In particular, Jex [Ref. 10] has applied this task to a two-axis tracking system in which the primary task consisted of the critical task. The separate secondary task consisted of the sub-critical task. He demonstrated that noticeable decreases in the value of λ_c achieved in the primary task were obtained with only slight increases in the secondary level of instability, and that the combined tasks could be used for workload research.

B. PSYCHOPHYSIOLOGICAL MEASUREMENT

"Psychophysiological" is a combination of the adjectives psychological and physiological and refers to the inseparable

connection of the mental and neuromuscular responses of the human operator. Since these responses are considered to be directly related to the operator's state of stress (workload), they have been measured quite extensively as both monitors and indicators of workload. Analyses of such psychophysiological responses as heart and respiration rates, galvanic skin response, electrocardiograms and electroencephalograms have been performed.

Spyker et al. [Ref. 4] performed a complex analysis of sixty-four such responses during a two-axis tracking task. Using a considerable amount of interface and data processing equipment he arrived at a reliable workload index consisting of ten of these responses.

Kahneman [Ref. 11] recorded and compared just three responses, including the somewhat unusual measurement of changes in the subject's pupil size, during a paced mental task. He observed that this pupil size variation provided the most consistent results.

The technique of pupillometry, or the method of measuring changes in pupil size, has primarily remained in the realm of the behavioral scientist. The preliminary work of Hess and Polt [Ref. 12] initiated a surge of research in this area which resulted in the acclamation of pupillometry as a very effective indicator of the human condition. It has been shown to indicate an individual's attitudes and preferences [Ref. 13], prejudices [Ref. 14], sexual arousal [Ref. 15], even latent homosexual tendencies [Ref. 16]. Pupillary

variations have been used to indicate psychiatric disorders [Ref. 17], as well as physical ills [Ref. 18]. Hess [Ref. 19] provides an excellent assessment and summary of the more important research performed in these areas.

Cognitive demand experiments, i.e., those requiring mental effort, have also been performed and related pupil changes to levels of mental activity. Hope [Ref. 20] showed that increasing levels of difficulty in a mental arithmetic problem resulted in proportionate pupil dilations and noted the phenomena of pupillary constriction when the difficulty level was excessive, causing incorrect replies. He referred to this as a mental "overload."

Payne et al. [Ref. 21] reported that changes in pupil diameter were a sensitive and reliable indicator of information processing capacity and were related to levels of workload.

Edwards [Ref. 22] showed similar correlation with information processing and reported the same pupil constriction for mental overloads as shown by Hope [20].

III. EXPERIMENT

A. APPLICATION OF THEORY

1. Tracking Tasks

a. Cross-Adaptive

Two variations of the critical task[1] were used as the primary and secondary tasks in a cross-adaptive system. The sub-critical task served as the controlled element of the primary task, and a variable sub-critical (i.e., where λ varies continuously at levels below λ_c) was used as the secondary (loading) task. This was done for several reasons:

(1) The critical task is simple, easily mechanized and can be used on a small analog computer.

(2) The level of instability (λ) is directly related to the level of difficulty of the task and can be used both to set the level of difficulty for the primary task (the basis for workload comparison) and to vary the level in the secondary task (thus becoming the adaptive variable).

(3) The tasks required no input. The random operator remnant is sufficient to excite and maintain the unstable tasks.

---

[1]Although the term "critical" specifically refers to the critical instability task in which λ is monotonically increased until reaching λ_c, it has been common to refer to the general (i.e., first-order divergent) compensatory tracking task from which λ_c is obtained as the critical task. Unless otherwise noted, the author has used the term in this general sense.

(4) In addition to its function as the adaptive variable, the secondary instability becomes the indirect measure of primary task workload.

The performance measure employed, although not as sophisticated as desired, was the deviation of the absolute value of the primary task error (e_p) from a fixed criterion (e_f). The criterion was established by determining the rms error of the principal subject during tracking runs at a median level of primary task instability (λ_p) and was set at 0.5 cm. (of displayed error). Thus primary task performance (P_p) is defined:

$$P_p = e_f - |e_p| \qquad (2)$$

The adaptive logic used was similarly simple. The time rate of change of the adaptive variable ($\frac{d\lambda_s}{dt}$) was related to P_p in a manner similar to Kelley's approach [Ref. 7, p. 554] i.e.,

$$\frac{d\lambda_s}{dt} = K P_p \qquad (3)$$

where the gain (K) is set:

$$K = 0.4 \text{ rad./cm.-sec.}^2 \qquad (4)$$

To prevent excessive instability rates and to provide the operator with a greater opportunity to regain control, the rate was limited in the following manner:

$$\left.\frac{d\lambda_s}{dt}\right|_{max} = 0.14 \text{ rad./sec.}^2, \quad P_p \geq 0.35 \qquad (5)$$

$$\left.\frac{d\lambda_s}{dt}\right|_{max} = -0.26 \text{ rad./sec.}^2,\ P_p \leq -0.65 \qquad (6)$$

Thus when the operator exceeded the criterion, the secondary task instability rate decreased almost twice as fast as it increased when he tracked within the criterion.

A further limit was placed on the instability to prevent the secondary task from ever becoming stable which, in this particular mechanization, would cause control reversal. Thus:

$$\lambda_s\big|_{min} = 0.16 \text{ rad./sec.} \qquad (7)$$

An initial value of instability (λ_{s_o}) was provided to insure that the operator was loaded in both axes at the beginning of the run. This relatively low initial value was:

$$\lambda_{s_o} = 1.0 \text{ rad./sec.} \qquad (8)$$

A diagram of the cross-adaptive system is shown in Fig. (2).

b. Single-Axis Sub-Critical

In order to provide a common ground for comparing the two different methods of workload measurement, the primary task alone had to be performed for pupillometric measurements. By disconnecting both the secondary task and the cross-adaptive network, the single-axis sub-critical task was generated.

c. Single-Axis Critical

The critical task[2] was used to determine each subject's tracking skill and level of proficiency during the experiment. It was used to normalize measured instability levels obtained during the tracking tasks and to indicate the level of total control capacity.

Although an automatic, rate-switched mechanism (autopacer) was recommended by Jex et al. [9], this was not possible in this experimental set-up and the value of λ_c had to be obtained by manually increasing the level of instability on the task from zero to the point where control was lost.

2. Pupil Diameter

The results obtained by Kahneman [11] and by Anderson [Ref. 23] led to the selection of pupil diameter changes as the psychophysiological indicator of workload. The authors indicated a correlation between level of difficulty and pupil size variation and utilized "information processing capacity" in a manner analogous to the use of "reserve capacity" in this study.

B. PROCEDURE

Three subjects[3], chosen on the basis of motivation, availability and previous tracking or flying experience,

---

[2] Now used in its specific sense (i.e., $\lambda \to \lambda_c$).

[3] Several other subjects were considered to provide a larger and more meaningful data base. After several trial runs they had to be rejected due to eye characteristics (i.e., little contrast between iris and pupil or partial covering of the pupil by the eyelid) which caused the pupil measuring equipment to give erroneous readings. A similar difficulty was reported by Burns [24].

performed the critical tracking tasks described previously. The tasks were mechanized on a small analog computer according to the circuit diagram shown in Fig. (14). Tracking errors were displayed as the displacement of a dot from the center of a cathode ray tube (CRT) located 15.5 in. (39.4 cm.) in front of the subject's eye.

Control was affected with an isometric (force) stick. The controller was mounted on a tablet arm to provide forearm support and positioned slightly ahead and to the right of the subjects, who were all right-handed. To reduce operator fatigue and to allow higher response frequencies, the sensitivity of the controller was set at a fairly high level. The ratio of display deflection to control force was 2.84 in./lb. (1.62 cm./N).

Continuous pupil diameter measurements were made using a Space Sciences Model 831 Television Pupillometer. This device uses a closed-circuit television system to observe the eye and a signal processing unit to measure and display pupil diameter. The eye was illuminated by a small, near-infrared, low-intensity light source which was mounted to one side of the adjustable camera. The subject's head was positioned in a brace consisting of a chin support and a forehead restraint (Fig. 3). An integral two-channel strip chart recorder in the processing unit recorded pupil diameter in a range from zero to ten millimeters. The second channel was used to record either secondary task instability (λ_s) during the cross-adaptive task, or displayed error (e_p)

during the sub-critical tracking runs. An additional recorder was also used to record total displayed error ($e_T = |e_p| + |e_s|$) during the cross-adaptive runs. A picture of the experimental set-up is shown in Fig. (4).

Seated before the display, each subject positioned his head comfortably in the restraint while focusing on the center of the elevated CRT. The display was raised 14.5 degrees above eye level to force the subject to look up, thus providing better illumination of the eye and removing any interference between the eyelid and pupil.

The experiments were performed in an acoustically protected chamber with constant, dim back lighting. The chamber was located in the Human Engineering Laboratory of the Naval Postgraduate School.

The experiments were divided into three phases, each corresponding to the different tracking tasks to be performed. During the initial phase the subject was given a warm-up period to become familiar with the task and to achieve a relatively stable performance level. He was then asked to perform the cross-adaptive critical task, and reminded that control of the primary task was paramount. This meant that he should try to keep the vertical position of the dot (indicating e_p) in the center of the CRT as much as possible or to center it as quickly as possible when displaced. He was requested to keep the horizontal position of the dot (e_s) at least within five cm. of the center.

A series of runs, each lasting from 90 to 150 seconds, were then performed at various levels of primary task instability. Table Ia. indicates the number of runs performed by the subjects at each of these levels.

Recordings of secondary task instability and total error were made during each run. Although this experimental phase was primarily aimed at obtaining values of λ_s as a workload indicator, pupillary reactions were recorded to determine if any correlation existed between pupil dilation and λ_s.

In the second experimental phase, subjects performed the critical instability task in the vertical axis. Each subject performed the task five times with run lengths averaging 20 to 30 seconds. Values of λ_c were recorded along with the subsidiary measurement of pupil diameter. Figure (5) shows a typical time history of the critical task instability levels and pupillary variations.

The final phase of the experiment consisted of having the subjects perform a series of sub-critical single-axis tracking tasks at the same instability levels (λ_p) as used in the cross-adaptive tasks. Measurements of pupil diameter and of displayed error (e_p) were recorded. The number of runs per subject at each level of primary instability is shown in Table Ib.

C. DATA ANALYSIS

To reduce the effects of inter-subject variation, the recorded data was normalized. Instability levels were

represented as a percentage of the subject's mean critical value of instability ($\overline{\lambda}_c$), i.e.,

$$\lambda^*_{p,s} = \frac{\lambda_{p,s}}{\overline{\lambda}_c} \times 100 \qquad (9)$$

A procedure recommended by Pratt [Ref. 25] and one commonly used in pupillometric measurements was to express the pupil diameter (D) as a percentage change from the average unloaded initial diameter ($\overline{D_o}$), i.e.,

$$D^* = \frac{D - \overline{D_o}}{D_o} \times 100 \qquad (10)$$

The determination and averaging of this initial diameter, however, was far from simple. Figure (9) shows the wide and random variation of one subject's pupil diameter during two periods of rest following consecutive tracking runs. These effects are caused both by a small amount of random noise introduced in the optical neuromuscular system (akin to McRuer's "remnant" [8]) as described by Stanten and Stark [Ref. 26] and by operator cognitive processes (i.e., day-dreaming). The latter effect is presumed to be predominant and is eliminated when the operator is loaded and forced to pay full attention to the tracking task.

The figure also indicates that there is a significant difference between the mean values. Thus a determination of $\overline{D_o}$ is required for each separate run.

In order to preclude an inordinate amount of tracking and data reduction time, this type analysis (i.e., numerically

averaging pupil diameter every three seconds during an initial 90 second period) was not performed. Instead, a 15-30 second rest period before each run was used as the baseline and the pupil recordings were visually averaged to obtain $\overline{D_o}$.

Analysis of pupil measurements recorded during both the cross-adaptive and sub-critical tracking tasks required similar averaging techniques. Although not elegant from an engineering viewpoint, this method provided adequate and representative mean data. It may have been partially responsible, however, for introducing scatter in the data.

During the analysis of the instability data, a consistently recurring phenomena was observed. Peak values of secondary instability (λ_{sm}), which indicated the highest level of instability achieved before the primary performance rapidly deteriorated, were noted. The pupil diameter recording also showed a similar peak (D_{sm}) within a few seconds. These values were measured and recorded. Figure (7) indicates this phenomena quite clearly.

Values of asymptotic (steady state) secondary instability (λ_{ss}) were also visually averaged over the observed time intervals (Δt_s) during which they were maintained. Pupil diameter measurements (D_{ss}) were also averaged during this interval, and the values for each were recorded.

Figures (6) to (8) show typical time histories of the variables for the three tracking tasks.

All data was normalized and averaged for each subject at each level of λ_p^*. Except for steady state values, the

mean and standard deviation was calculated using standard statistical formulae, i.e.,

$$\bar{x} = \frac{\sum_{i=1}^{N} x}{N} \tag{11}$$

and,

$$\sigma = \left[\frac{\sum_{i=1}^{N} x^2 - \left(\sum_{i=1}^{N} x \right)^2 / N}{N - 1} \right]^{1/2} \tag{12}$$

Where,

$x = \lambda_c, \lambda^*_{sm}, D^*, D^*_{sm}$

N = number of runs

The steady state values were weighted by the percentage ratio of the time interval (Δt_s) to the total run time (t_T), i.e.,

$$w = \frac{\Delta t_s}{t_T} \times 100 \tag{13}$$

Weighted averages and standard deviations were computed using the relations:

$$\overline{y_w} = \frac{\sum_{i=1}^{N} w_i y_i}{\sum_{i=1}^{N} w_i} \tag{14}$$

and,

$$\sigma_w = \left[\frac{\sum_{i=1}^{N} w_i y_i^2 - \left(\sum_{i=1}^{N} w_i y_i \right)^2 / N}{N - 1} \right]^{1/2} \qquad (15)$$

Where,

$$y = \lambda^*_{ss},\ D^*_{ss}$$

The averaged, normalized data obtained in this manner are shown in Tables II to IV.

IV. RESULTS

The principal results of this experiment are presented in Figs. (10) and (11). Using the cross-adaptive critical tracking task as an indicator of workload, Fig. (10) shows the overall decrease in the steady state values of normalized secondary instability with increasing normalized primary instability. The demonstrated relationship between secondary instability and primary workload thus validates the cross-adaptive technique. Simply stated, this shows that as the primary task becomes more difficult, the level of difficulty maintained in the secondary task becomes less, indicating the previously predicted increase in primary task workload. Thus, the task difficulty is reflected by its workoad and vice versa.

The results of using changes in pupil diameter as a measure of workload are shown in Fig. (11)[4]. The proportional relationship between level of difficulty and increase in pupil diameter is evident and thus reflects the increasing workload. An apparent increase in the rate of change of pupil size is noted at approximately 50 percent of the total

---

[4]Data are presented in this figure for only two subjects. The third subject became rapidly fatigued during this experimental phase. This fatigue caused his eyelids to droop slightly which in turn interfered with the correct measurement of his pupil diameter.

operator control capacity and indicates, perhaps, a mental "shifting of gears" at the midpoint of this capacity.

Due to the large variance in the data, no regression analyses were performed. The mean trends, however, are clearly evident and support the theoretical hypothesis.

Additional results obtained from the experiment are shown in Figs. (12) and (13). The variance of peak values of normalized primary instability is shown in Fig. (12). The same trend as shown previously in Fig. (10) is evident. This parameter could thus be used as an additional workload index.

Fig. (13) shows the results of the critical task measurements. It can be seen that two subjects dilated at fairly constant rates until reaching approximately 90 percent of critical, at which time the dilation rate greatly increased until reaching critical. The unexpected phenomenon of initial pupillary constriction with increasing instability is observed in one subject, and can be explained in light of his admitted tracking strategy.

This subject had previously performed a large number of critical instability tracking tasks and had developed a strategy which enabled him to achieve consistently high values of λ_c. By consciously relaxing during the early phases of the run when tracking is relatively simple, he would store control energy for the highly unstable regions (essentially switching from a low to a high gain). The pupillometer recorded this reduction in stress (workload) by either an accurate measurement of actual pupil constriction

or by the inaccurate measure of the pupil partially blocked by a lowered (restful) eyelid. Dilation occurred when the instability loaded the subject beyond his capacity to relax and is shown to have a similarly high rate near critical.

One additonal result was obtained with the critical task. As indicated in Fig. (5), rapid pupil constriction following loss of control is evident. This is a manifestation of the "overload" condition of Hope [20] and Edwards [22].

The additional measurements of changes in pupil diameter during the cross-adaptive task at corresponding levels of steady state and peak secondary instability, provided no discernable information. It is assumed that the complex interaction of primary and secondary tracking tasks as seen through the eye, made analysis of pupillary responses fruitless.

The results obtained from the cross-adaptive critical tracking task have been compared to those of Jex et al. [Ref. 27]. Although not specifically employed to indicate operator workload, their more sophisticated cross-adaptive task, coupled with measurements of operator describing functions, provided three results similar to those of this experiment:

1. Three primary task controlled elements of increasing difficulty resulted in decreasing levels of cross-coupled instability.

2. Small increases in effective time delay and remnant were measured.

3. Operator behavior (tracking strategy) in the primary task as determined by describing function measurements was not changed by secondary task loading, indicating that he had not changed his primary task performance criterion.

V. CONCLUSIONS AND RECOMMENDATIONS

The following conclusions concerning this research are drawn:

1. This thesis has been successful in mechanizing two techniques for determining the level of workload of an operator controlling a manual control system.

2. The techniques of using the cross-adaptive critical tracking task and of measuring changes in pupil size provide valid indices of operator workload.

3. The cross-adaptive task can be mechanized on a small computer using simple performance measurements and logic which will yield results comparable to those obtained from larger, more complex systems.

4. An additional workload index, the peak value of secondary task instability, is available and can be used to confirm results obtained from the two principal indices.

5. The measurement of critical instability levels provides a good measure of an operator's total control capacity and enables the workload index to be expressed concretely.

6. Pupil measurements are highly inter- and intra-subject variant. Additional variance may have been introduced by a relatively crude method of data reduction.

7. The simultaneous measurement of cross-adaptive instability and changes in pupil diameter provides no meaningful data.

The following recommendations are offered:

1. The cross-adaptive tracking task should be coupled with the recently developed hybrid describing function analyzer [Ref. 28] as a more effective measure of workload.

2. The use of a smoothing filter and of similar performance measurement techniques as presented in Ref. 27 should be incorporated in the cross-adaptive mechanization.

3. A more realistic and less fatiguing display format should be used (e.g., a simulated artificial horizon).

4. A method of interfacing pupillometric output with an off-line data processor including programs for data averaging should be found. This would make pupillometry more appealing as a workload measure for tracking tasks.

APPENDIX A

TABLES

TABLE I

Number of Experimental Runs per Subject

a.
Cross-Adaptive Tracking Task

	Primary Task Instability (λ_p)						
Subject	1.0	1.5	2.0	2.5	3.0	3.5	4.0
RAH	-	-	6	11	15	8	5
HGC	-	-	4	3	3	-	-
CMB	5	5	5	-	5	-	-

b.
Sub-Critical Tracking Task

	Primary Task Instability (λ_p)						
Subject	1.0	1.5	2.0	2.5	3.0	3.5	4.0
HGC	-	-	4	3	3	-	3
CMB	3	3	3	3	3	-	3

TABLE II

Normalized Instability Levels and Pupil Diameters for Sub-Critical Tracking Task

Subject	Normalized Primary Instability (λ_p*)	Normalized Pupil Diameter (D*) (mean / std. dev.)
HGC	32.34	7.85 (2.60)
	40.43	12.60 (0.70)
	48.51	12.38 (1.09)
	64.70	18.12 (5.42)
CMB	21.93	3.53 (9.50)
	32.89	5.52 (2.17)
	43.86	10.94 (2.58)
	65.79	35.79 (22.86)
	87.70	50.03 (5.68)

TABLE III

Normalized Instability Levels
for Cross-Adaptive Critical Tracking Task

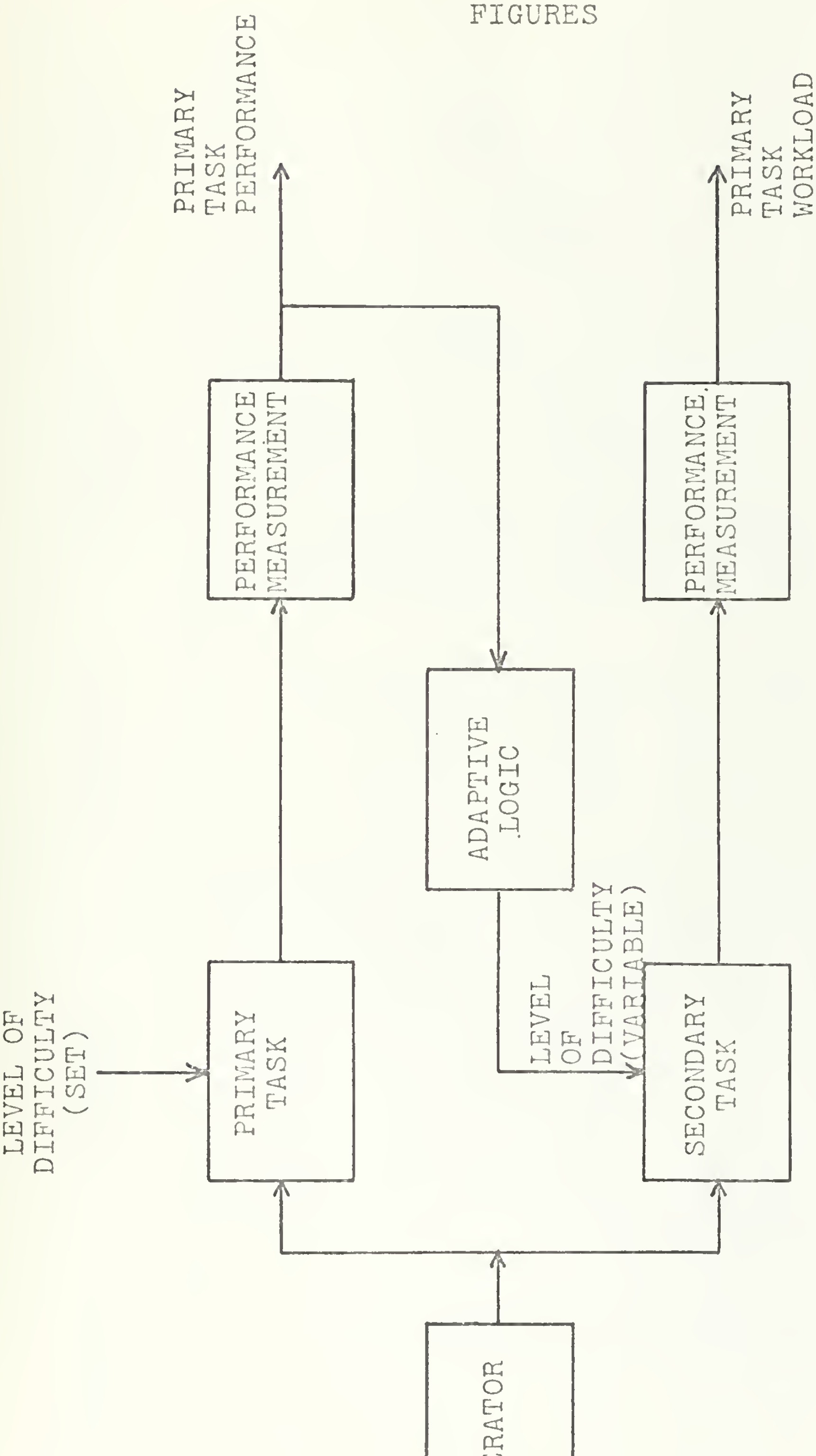
Subj.	Normalized Primary Instability (λ_p*)	Normalized Steady State Secondary Instability (λ^*_{ss})	Normalized Peak Secondary Instability (λ^*_{sm})
		(mean / standard deviation)	
RAH	31.64	65.06 (9.93)	74.19 (10.15)
	39.56	65.03 (4.59)	76.58 (6.41)
	47.47	62.75 (11.06)	75.68 (12.97)
	55.38	43.16 (13.47)	62.18 (19.45)
	63.29	30.66 (15.44)	35.24 (19.46)
HGC	32.32	58.02 (11.23)	55.63 (10.58)
	40.43	68.20 (3.00)	56.89 (6.58)
	48.51	42.65 (8.22)	55.63 (10.58)
CMB	21.93	79.32 (13.58)	90.34 (9.34)
	32.89	62.79 (9.39)	75.29 (9.80)
	43.86	27.43 (21.10)	42.32 (22.9)
	65.79	7.04 (3.90)	29.0 (14.17)

TABLE IV

Normalized Pupil Diameters
for Cross-Adaptive Critical Tracking Task

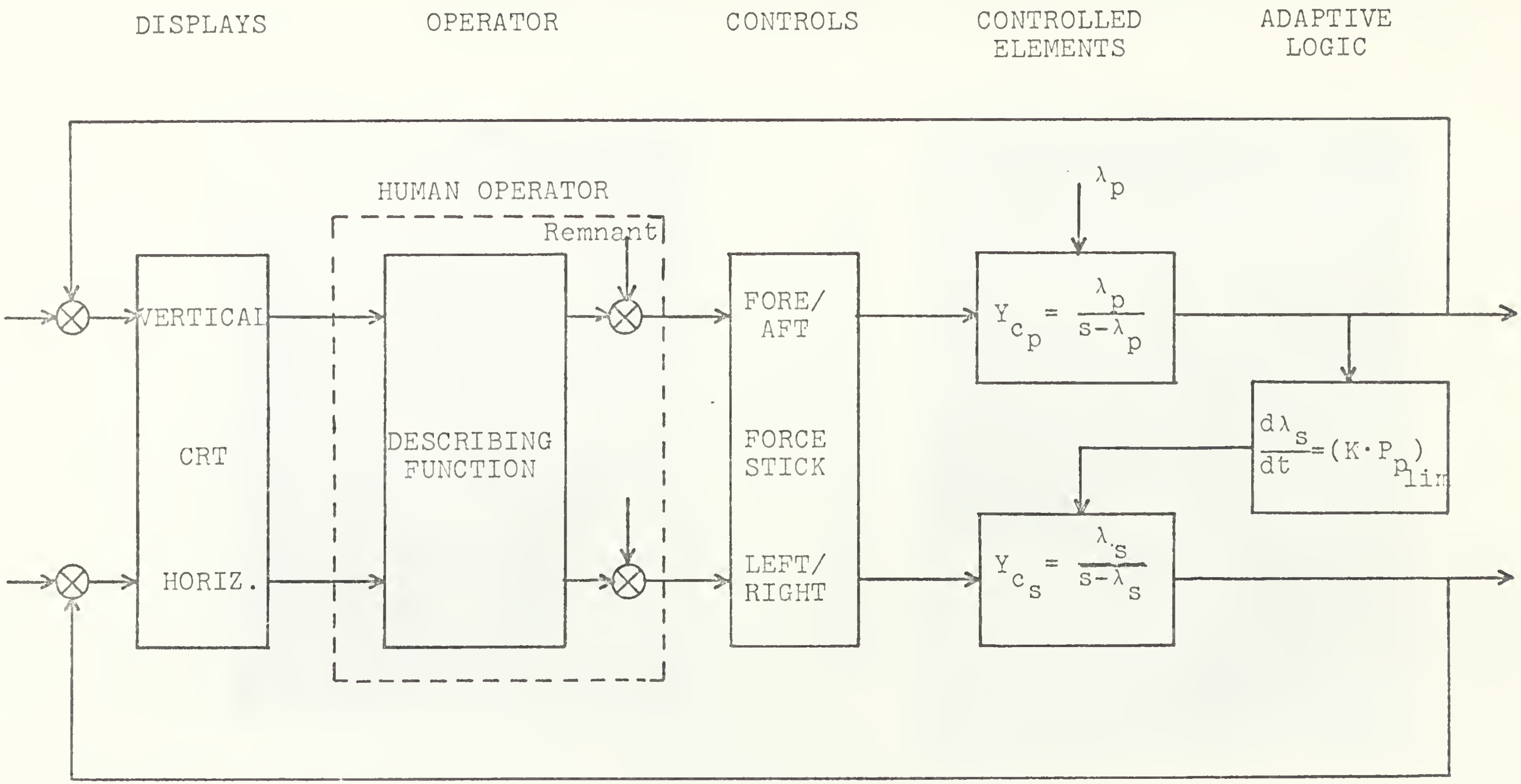
Subj.	Normalized Primary Instability (λ_p*)	Normalized Steady State Pupil Diameter (D^*_{ss})	Normalized Peak Pupil Diameter (D^*_{sm})
RAH	31.64	4.63 (5.13)	35.95 (24.99)
	39.56	2.87 (4.19)	18.87 (12.75)
	47.47	0.82 (5.53)	14.35 (8.86)
	55.38	-5.35 (7.98)	2.69 (6.78)
	63.29	-2.56 (8.70)	12.36 (5.36)
HGC	32.34	18.66 (4.34)	18.23 (4.84)
	40.43	16.65 (4.60)	6.86 (1.51)
	48.51	3.03 (7.77)	18.04 (6.37)
CMB	21.93	15.06 (13.02)	36.45 (15.73)
	32.89	-4.12 (20.82)	19.50 (15.56)
	43.86	-6.35 (12.91)	6.86 (8.81)
	65.79	2.58 (17.47)	14.10 (15.56)

APPENDIX B
FIGURES



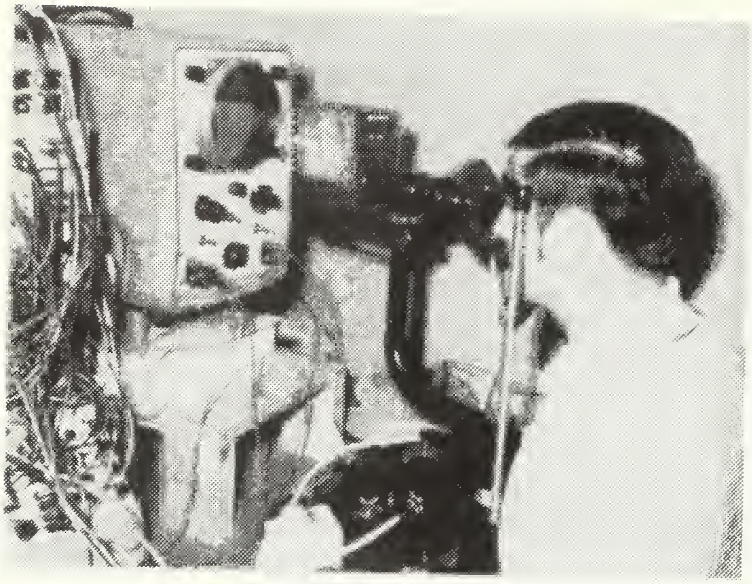
A Cross-Adaptive System

Figure (1)



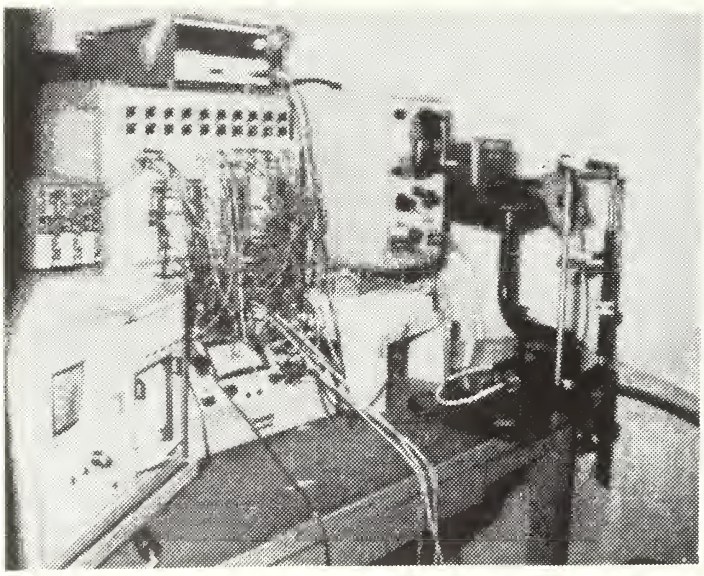
Block Diagram of the Cross-Adaptive Critical Task

Figure (2)



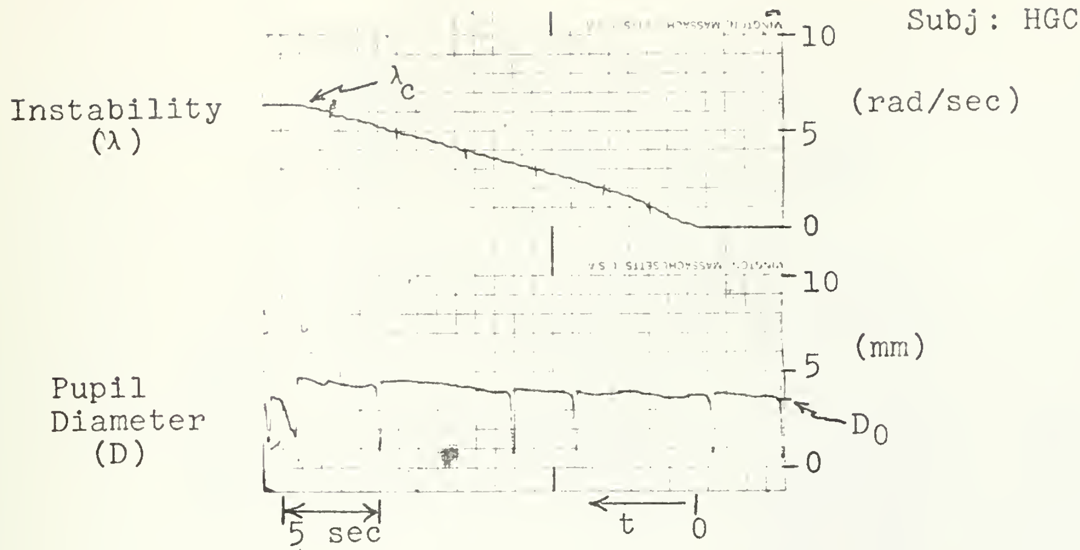
Close-Up of Subject Tracking
(indicating TV Pupillometer)

Figure (3)



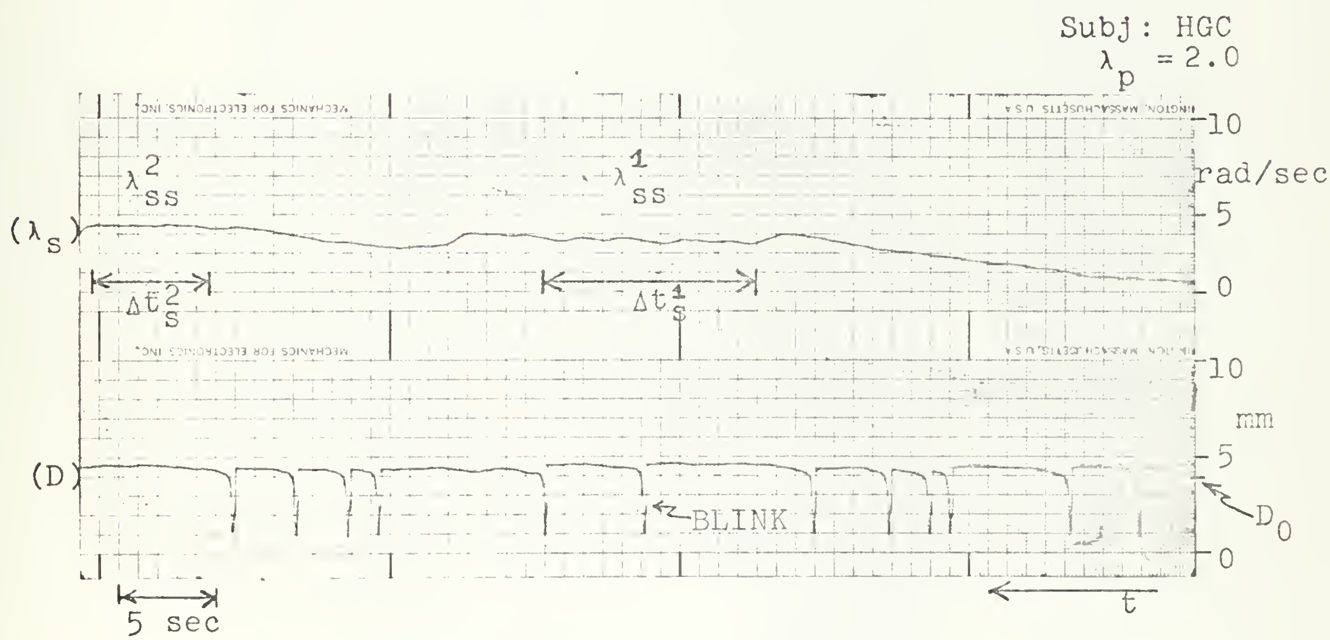
Experimental Set-Up

Figure (4)



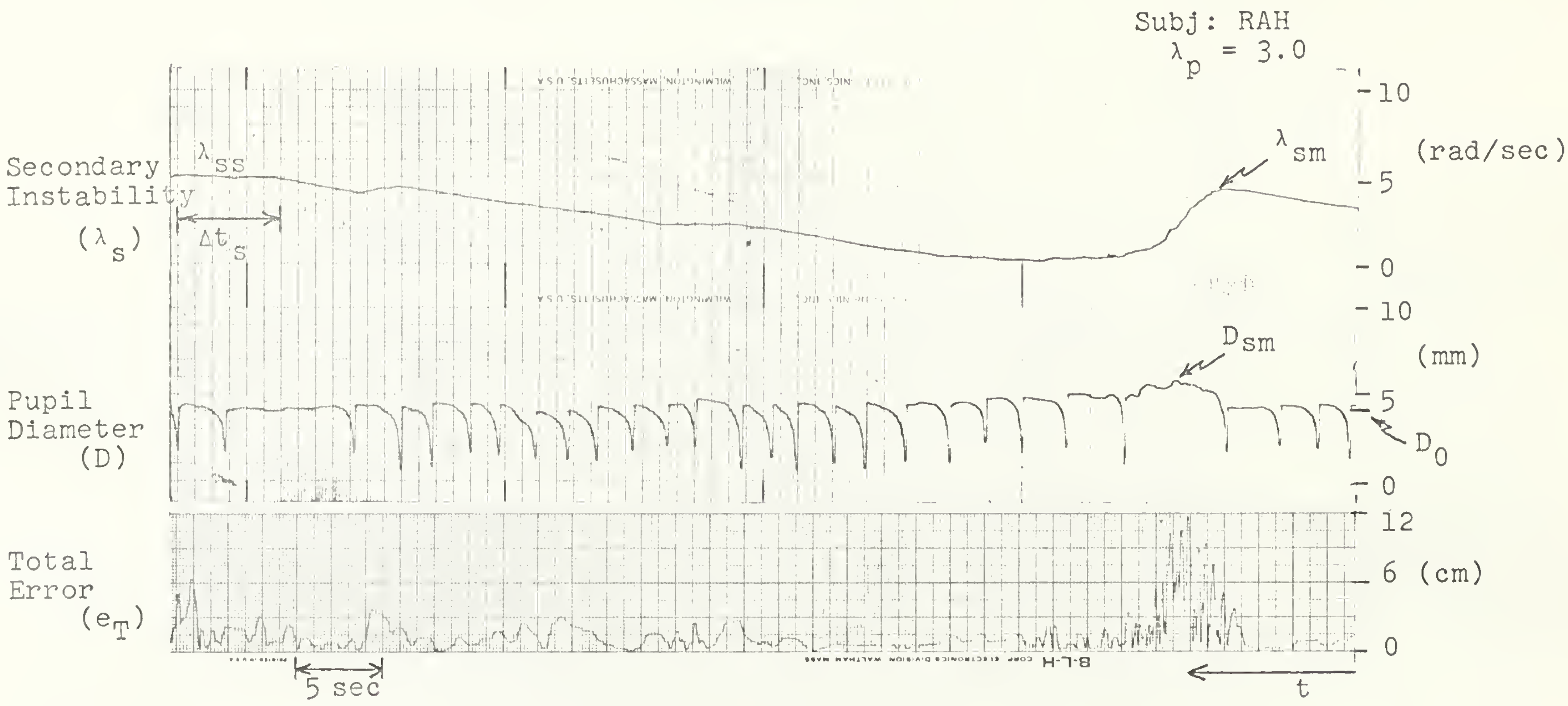
Time History of Typical Critical Instability Task

Figure (5)



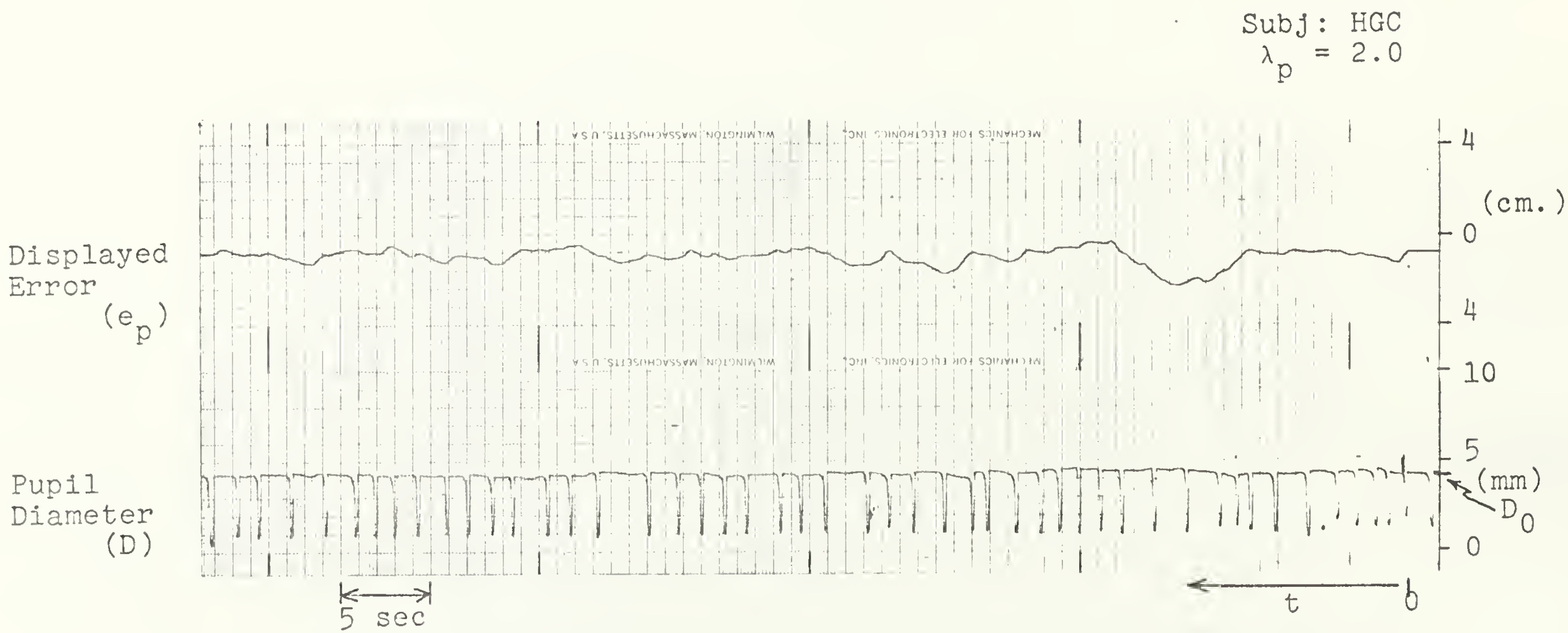
Time History of Typical Cross-Adaptive Critical Tracking Task
(indicating steady state values)

Figure (6)



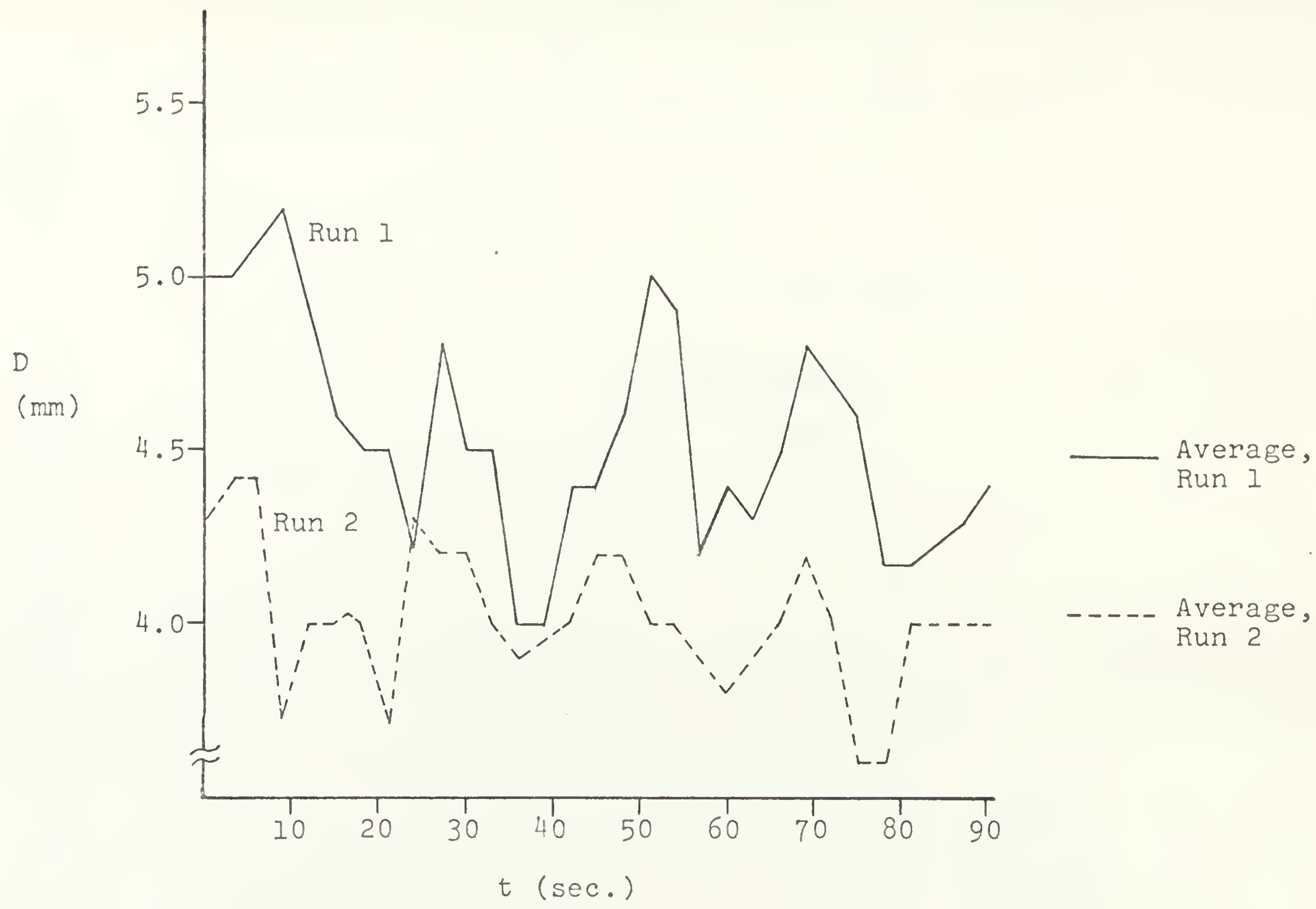
Time History of Typical Cross-Adaptive Critical Tracking Task
(indicating peak and steady state values)

Figure (7)



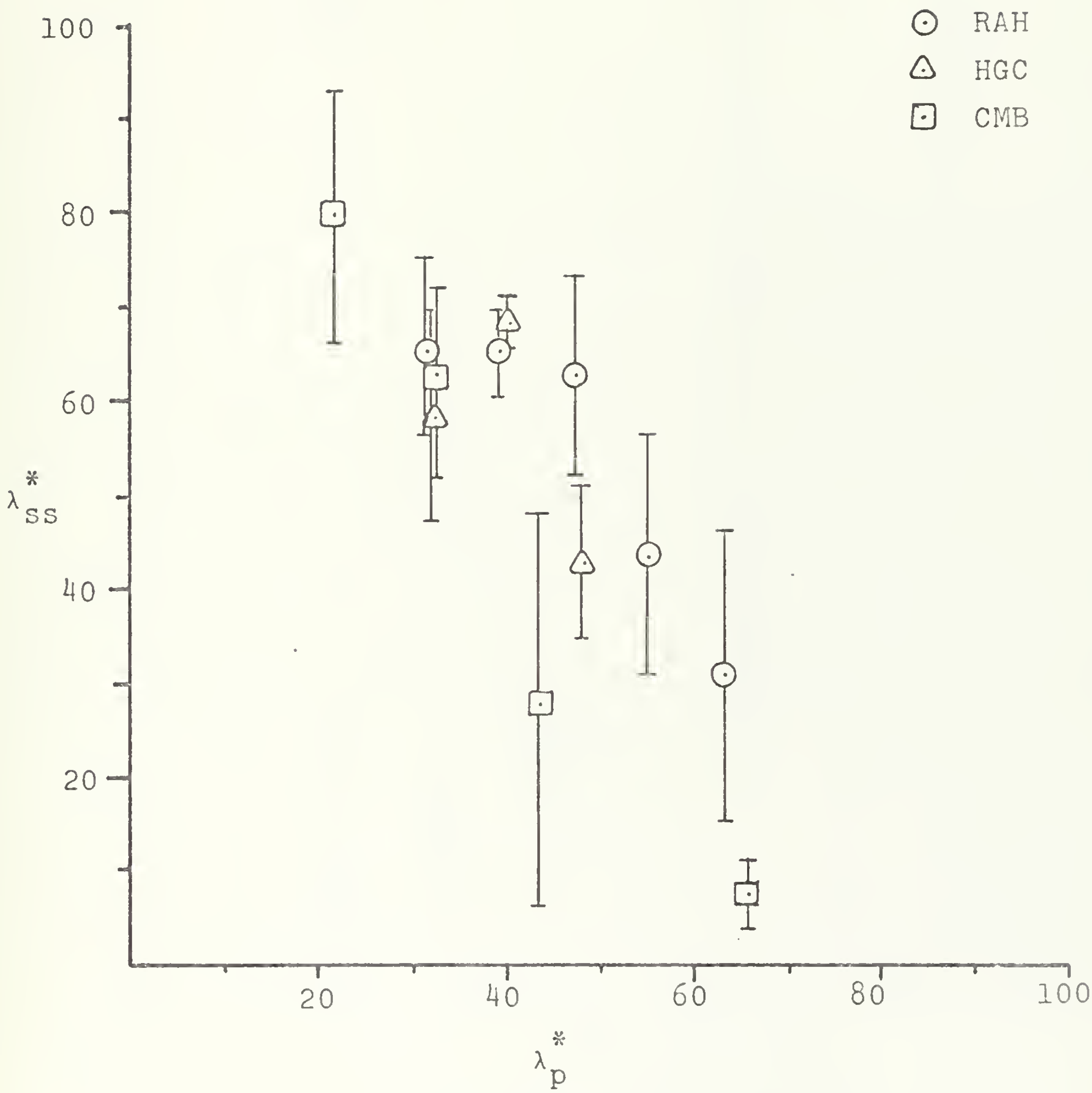
Time History of Typical Sub-Critical Tracking Task

Figure (8)



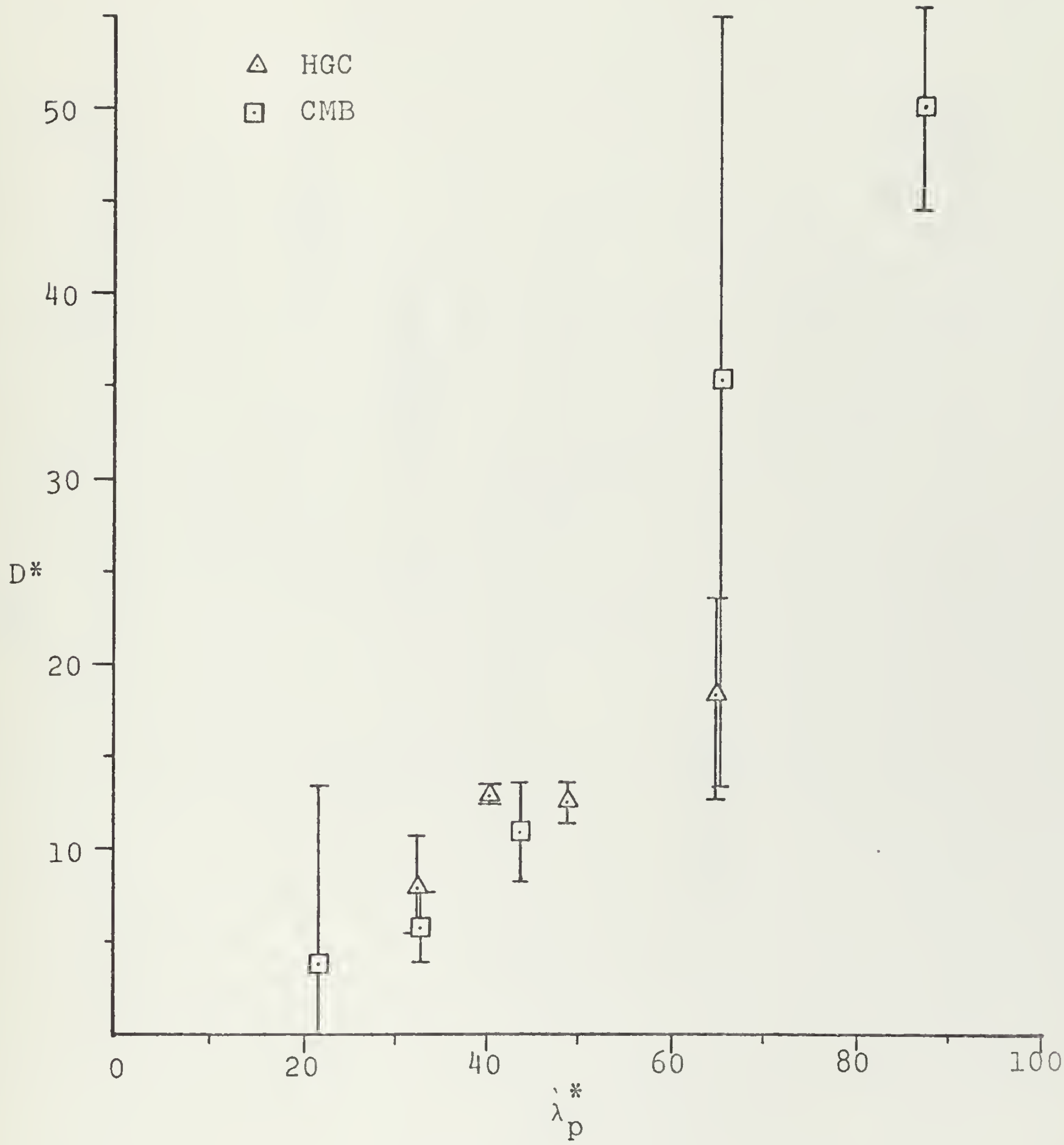
Time Histories of Typical Unloaded Pupil Diameter

Figure (9)



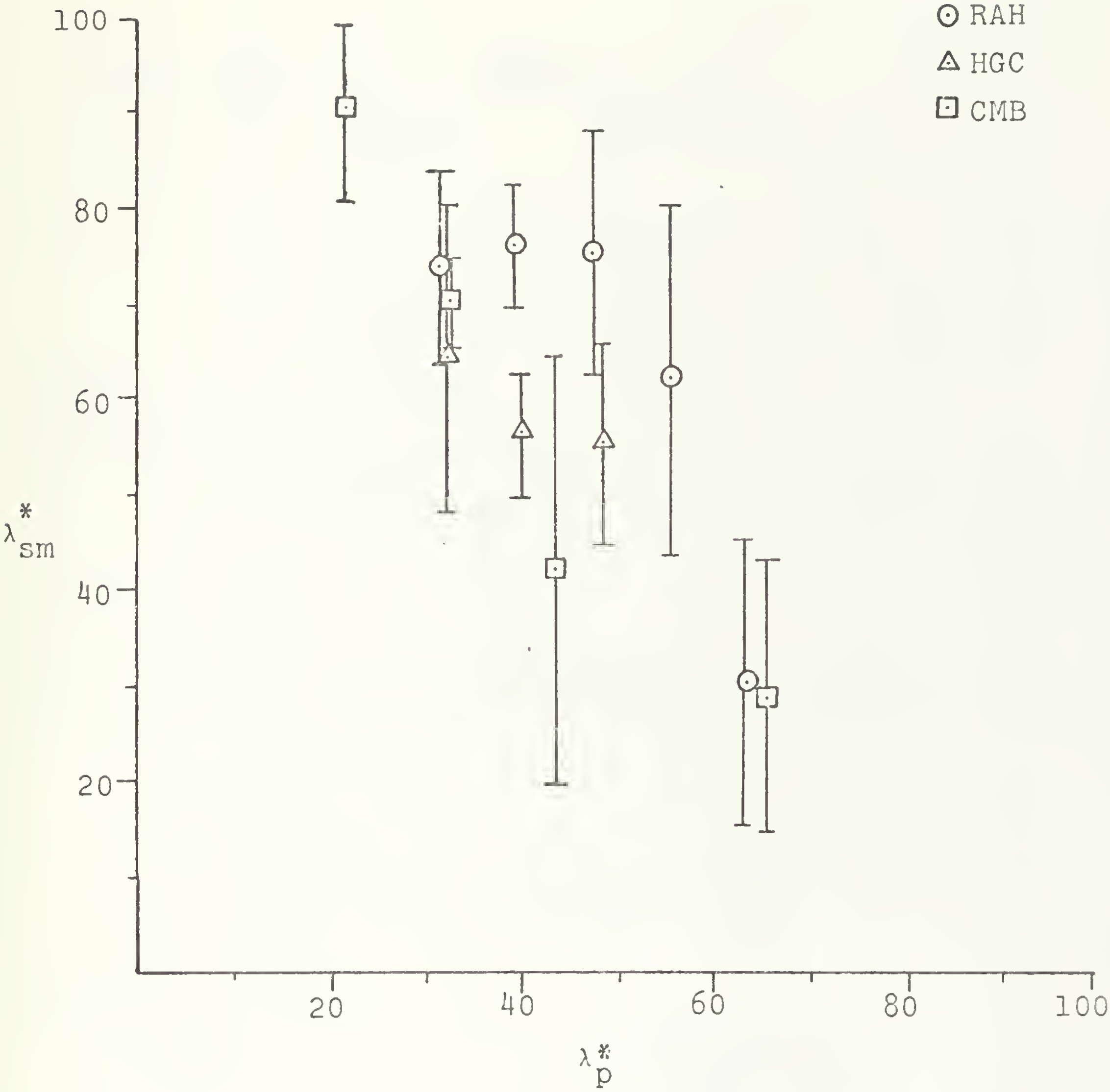
Normalized Steady State Secondary Instability vs. Normalized Primary Instability for Cross-Adaptive Task

Figure (10)



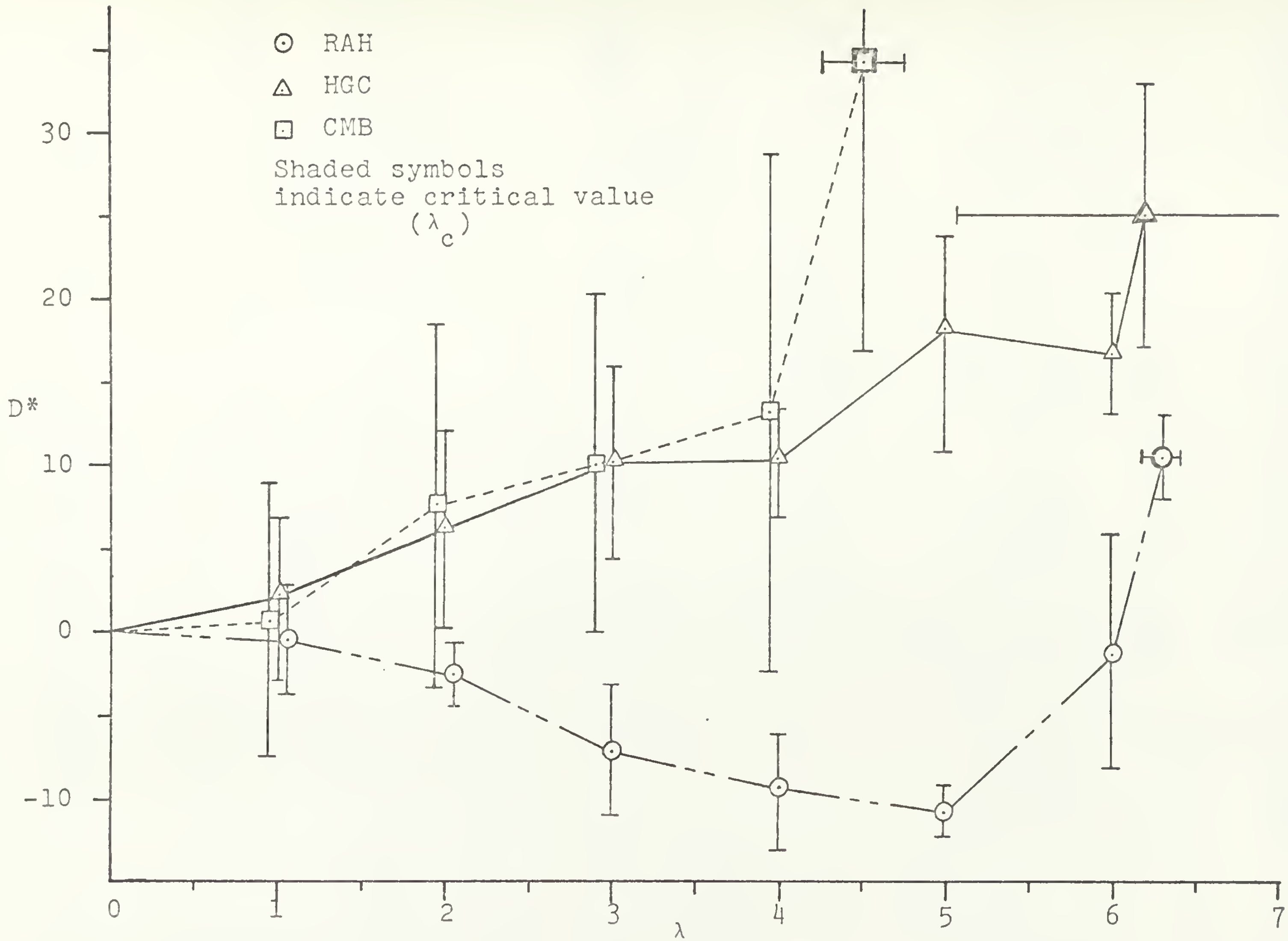
Normalized Pupil Diameter
vs.
Normalized Primary Instability
for
Sub-Critical Tracking Task

Figure (11)



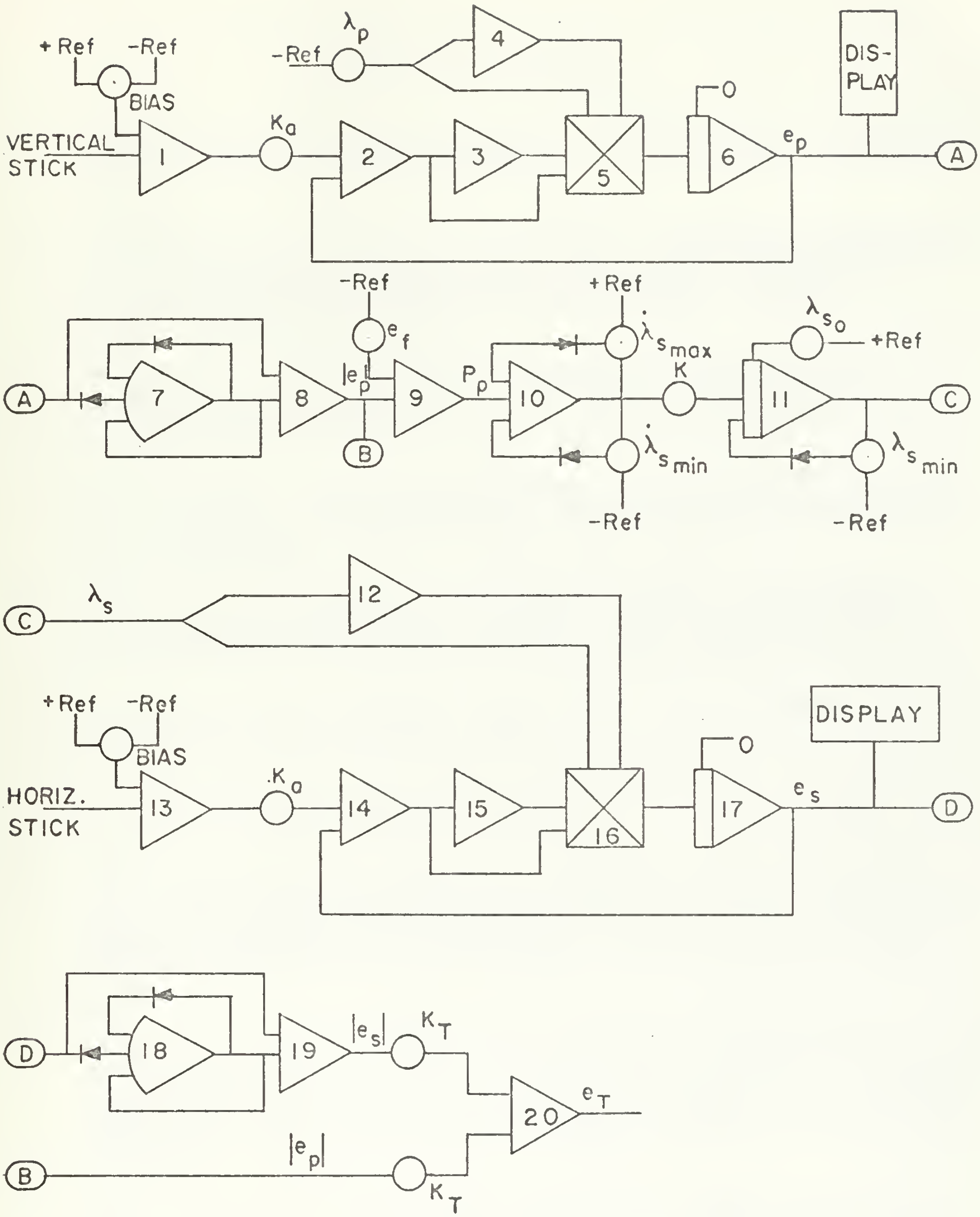
Normalized Peak Values of Secondary Instability
vs.
Normalized Primary Instability

Figure (12)



Normalized Pupil Diameter vs. Critical Task Instability Levels

Figure (13)



Circuit Diagram
For Cross-Adaptive Critical Tracking Task

Figure (14)

LIST OF REFERENCES

1. Kelley, C. R., and M. J. Wargo, "Cross-Adaptive Operator Loading Tasks," Human Factors, v. 9, (5), pp. 395-404, 1967.

2. Anderson, P. A., and M. L. Toivanen, Effects of Varying Autopilot Assistance and Workload on Pilot Performance in the Helicopter Formation Flight Mode, Janair Report 680610, Mar. 1970.

3. McDonnell, J. D., Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities, AFFDL TR-68-76, Dec. 1968.

4. Spyker, D. A., S. P. Stackhouse, A. S. Khalafalla, and R. C. McLane, Development of Techniques for Measuring Pilot Workload, NASA CR-1888, Nov. 1971.

5. Hess, R. A., The Use of a Nonadjectival, Nonordinal Linear Rating Scale in a Single Axis Tracking Task, U. S. Naval Postgraduate School, NPS-57He7191A, Sep. 71.

6. Knowles, W. B., "Operator Loading Tasks," Human Factors, v. 9, (5), pp. 155-161, 1963.

7. Kelley, C. R., "What is Adaptive Training," Human Factors, v. 11, (6), pp. 547-556, 1969.

8. McRuer, D. T., D. Graham, E. Krendel and W. Reisener, Jr., Human Pilot Dynamics in Compensatory Systems. Theory, Models and Experiments with Controlled Element and Forcing Function Variations, AFFDL TR-65-15, Jul. 1965.

9. Jex, H. R., J. D. McDonnell, and A. V. Phatak, A 'Critical' Tracking Task for Man-Machine Research Related to the Operator's Effective Delay Time. Part I: Theory and Experiments with a First-Order Divergent Controlled Element, NASA CR-616, Nov. 1966.

10. Jex, H. R., "Two Applications of a Critical-Instability Task to Secondary Workload Research," IEEE Trans. on Human Factors in Electronics, Dec. 1967, pp. 279-282.

11. Kahneman, D., "Pupillary, Heart Rate, and Skin Resistance Changes During a Mental Task," Journal of Experimental Psychology, v. 79, pp. 164-167, 1969.

12. Hess, E. H., and J. M. Polt, "Pupil Size as Related to Interest Value of Visual Stimuli," Science, v. 132, pp. 349-350, 1960.

13. Hess, E. H. and J. M. Polt, "Changes in Pupil Size as a Measure of Taste Difference," Perceptual and Motor Skills, v. 23, pp. 451-455, 1966.

14. Woodmansee, J. J., An Evaluation of Pupil Response as Measure of Attitude Toward Negroes, Progress Report, NSF. Grant GS-303, pp. 436-463. August, 1965.

15. Bernick, N., G. Borowitz, and A. Kling, Effect of Sexual Arousal and Anxiety on Pupil Size, Paper presented at Rocky Mountain Psychological Assn., Denver, Colo., May, 1968.

16. Hess, E. H., A. L. Seltzer, and J. M. Shlien, "Pupil Response of Hetero-and Homosexual Males to Pictures of Men and Women," Journal of Abnormal Psychology, v. 70, pp. 165-168, 1965.

17. Menninger, W. C., Pupillary Anomolies in Schizophrenia, Archives of Neurology and Psychiatry, v. 21, pp. 186-192, 1929.

18. Rubin, L. S., G. J. Barbero, and M. S. Sibinga, "Pupillary Dysfunction as a Concomitant of Cystic Fibrosis," Pediatrics, v. 38, pp. 865-873, 1966.

19. Hess, E. H., "Pupillometric Assessment," Research in Psychotherapy, v. 3, pp. 573-583, 1968.

20. Hope, R. E., Pupillary Size Change as Correlated with Mental Activity, M.S. Thesis, U. S. Naval Postgraduate School, 1971.

21. Payne, D. T., M. E. Parry, and S. J. Harasymiw, "Percentage of Pupillary Dilation as a Measure of Item Difficulty," Perception and Psychophysics, v. 4, pp. 139-143, 1968.

22. Edwards, B. R., Pupillary Size Changes During Information Processing Overload, M. S. Thesis, U. S. Naval Postgraduate School, 1972.

23. Anderson, R. O., and P. E. Pietrzak, "Pupil Dilation as a Measure of Workload," 3rd Annual NASA-Univ. Conf. on Manual Control, USC, Los Angeles, Calif., 1-3 Mar. 1967, pp. 305-308.

24. Burns, T. V., Pupil Diameter Variation in a Visual Interpretation Task, M.S. Thesis, U. S. Naval Postgraduate School, 1972.

25. Pratt, R. W., Cognitive Processing of Uncertainty: Its Effect on Pupillary Dilation and Preference Ratings, Ph.D. Thesis, U. of Illinois, 1969.

26. Stanten, S. F. and L. Stark, "A Statistical Analysis of Pupil Noise," IEEE Trans. on Bio-Medical Engineering, v. BME-13, (3), Jul., 1966.

27. Jex, H. R., W. F. Jewell, and R. W. Allen, Development of the Dual-Axis and Cross-Coupled Critical Tasks, Paper presented at the 8th Annual NASA-Univ. Conf. on Manual Control, U. of Michigan, Ann Harbor, Michigan, May 1972.

28. Warren, R. D., A Hybrid Computer Technique for Measuring Human Describing Functions and Remnant in Closed Loop Tracking Tasks, AeE. Thesis, U. S. Naval Postgraduate School, 1972.

<u>INITIAL DISTRIBUTION LIST</u>

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Chairman, Department of Aeronautics, Code 57Be Naval Postgraduate School Monterey, California 93940	1
4.	Assoc. Professor G. K. Poock, Code 55Pk Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
5.	Asst. Professor R. A. Hess, Code 57He Department of Aeronautics Naval Postgraduate School Monterey, California 93940	1
6.	Asst. Professor M. H. Redlin, Code 57Rp Department of Aeronautics Naval Postgraduate School Monterey, California 93940	1
7.	Mr. Duane T. McRuer Mr. R. Wade Allen Systems Technology Inc. 13766 S. Hawthorne Blvd. Hawthorne, California 90250	1
8.	LCDR Thomas E. McFeely Helicopter Anti-Submarine Squadron(Light) THIRTY Naval Air Station Lakehurst, New Jersey 08733	1

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Pupil Diameter and the Cross-Adaptive Critical Tracking Task; a Method of Workload Measurement

4. DESCRIPTIVE NOTES *(Type of report and, inclusive dates)*

Aeronautical Engineer's Thesis; June 1972

5. AUTHOR(S) *(First name, middle initial, last name)*

Thomas Edward McFeely

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1972	56	28

8a. CONTRACT OR GRANT NO. b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

Two new applications of established techniques for measuring an individual's level of stress (workload) in tracking tasks are presented. An indirect technique of measuring "reserve capacity" is utilized in a two-axis cross-coupled compensatory tracking task. A direct psychophysiological measurement is made by recording time histories of operator pupil diameter.

Results obtained indicate that each method yields a good index of workload, although considerable variance in the data is observed. The level of instability in the second axis of the cross-adaptive method is shown to be related to the level of workload in the primary axis. Increased pupil diameter is shown to be similarly related to operator workload. The simultaneous application of both techniques is determined to be inappropriate.

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
compensatory tracking						
human engineering						
human response						
manual control						
psychophysiological response						
pupillometry						
reserve capacity						
workload						

DD FORM 1 NOV 65 1473 (BACK)

S/N 0101-807-6821

Security Classification

A-31409